

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 10:19:03 ; Search time 1564.4 seconds

(without alignments)
11025.472 Million cell updates/sec

Title: US-09-922-895-2

Perfect score: 1065

Sequence: 1 ATGACACCTCTAGATAC.....CGAAGCTCTATGTGTTT 1065

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_estc3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	748.8	70.3	789	12	BG205056 RST24475
2	713	66.9	884	12	BG182330 RST1196 A
3	703.8	66.1	876	13	BI906283 603063222
4	651.6	61.2	745	12	BG204024 RST23417
5	474	44.3	907	9	AL552677 AL552677
6	437.2	41.1	934	12	BG460984 RST43688

7	432	40.6	1074	14	BM917063
8	410.8	38.6	881	12	BG460103
9	302.8	28.4	952	14	BM917763
10	294	27.6	747	12	BG220110
11	289.6	27.2	568	12	BF193051
12	279.6	26.3	1167	14	B0053936
13	268	25.2	533	12	BF193021
14	264.2	24.8	807	13	BI764263
15	218.8	20.5	639	14	W04836
16	211.4	19.8	659	10	B629533
17	210.6	19.8	487	10	B656336
18	209	19.6	804	13	BI68510
19	207	19.4	417	17	AQ341822
20	193.2	18.1	936	12	BF119806
21	190.8	17.9	581	10	B637546
22	187.8	17.6	307	17	A2800041
23	185.2	17.4	542	9	AA547303
24	184.8	17.4	2111	11	AK007808
25	184.6	17.3	672	10	BB638766
26	180.6	17.0	447	9	AA671573
27	176.6	16.6	835	13	BI821796
28	171.6	16.1	360	10	AA48515
29	169.4	15.9	867	13	BI106339
30	164.6	15.5	770	9	AJ452673
31	160.4	15.1	309	17	A2867933
32	156.6	14.7	312	12	BE756795
33	156	14.6	480	10	AM657263
34	155.4	14.6	863	13	BI661279
35	151.6	14.2	723	14	BM951933
36	151.4	14.2	551	12	BF081124
37	150.4	14.1	3005	11	AK019478
38	148	13.9	669	13	BI93893
39	147.2	13.8	885	13	BI082667
40	144.4	13.6	780	13	BI908127
41	143.4	13.5	965	14	BQ244555
42	142.8	13.4	452	9	A1851510
43	140.2	13.2	646	10	BB615654
44	140.2	13.2	665	10	BB628567
45	140.2	13.2	787	9	AU080004

ALIGNMENTS

RESULT 1	BG205056	789 bp	mrna	linear	EST 21-APR-2001
LOCUS	RST24475	Athersys RAGE Library	Homo sapiens	CDNA, mRNA sequence.	
DEFINITION	BG205056				
ACCESSION	BG205056.1	GI:13726743			
VERSION					
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1 (bases 1 to 789)				
	Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cochran, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.				
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)				
MEDLINE	21227151				
COMMENT	Contact: Scott J. Cain				
	Athersys, Inc.				
	3201 Carnegie Ave. Cleveland, OH 44115, USA				
	Tel: 216 431 9900				
	Fax: 216 361 9596				
	Email: scaine@atersys.com				
FEATURES	High quality sequence stop: 521.				
	Location/Qualifiers				

Query Match	Best Local Match	Similarity	Score	DB	Length
Match 764: Conservative	764: Conservative	98.6%	748.8	DB 12	769
		0: Mismatches	10: Indels	1: Gaps	
1	ATGACAACTCCTACTGATACAGTATGAGACCTTTGGTACACATCCTACTATGATGACGTG	16	ATGACAACTCCTACTGATACAGTATGAGACCTTTGGTACACATCCTACTATGATGACGTG	75	
61	GGCCTGCTCTGTGAAAAGCTGATACAGAGACAGTATGAGCCAGTTGTGGCCCGCTG	120	GGCCTGCTCTGTGAAAAGCTGATACAGAGACAGTATGAGCCAGTTGTGGCCCGCTG	135	
76	GGCCTGCTCTGTGAAAAGCTGATACAGAGACAGTATGAGCCAGTTGTGGCCCGCTG	135	GGCCTGCTCTGTGAAAAGCTGATACAGAGACAGTATGAGCCAGTTGTGGCCCGCTG	135	
121	TACTCCCGGTGTTACATGATGAGGCGCTTGGCCATGTTGGTGGTATGATGATCCTATA	180	TACTCCCGGTGTTACATGATGAGGCGCTTGGCCATGTTGGTGGTATGATGATCCTATA	180	
136	TACTCCCGGTGTTACATGATGAGGCGCTTGGCCATGTTGGTGGTATGATGATCCTATA	195	TACTCCCGGTGTTACATGATGAGGCGCTTGGCCATGTTGGTGGTATGATGATCCTATA	195	
181	AAATACAGAGGCTCCGAATATGACCAACATCTACCTGCTACACCTGAGCCATTTGGAC	240	AAATACAGAGGCTCCGAATATGACCAACATCTACCTGCTACACCTGAGCCATTTGGAC	240	
196	AAATACAGAGGCTCCGAATATGACCAACATCTACCTGCTACACCTGAGCCATTTGGAC	255	AAATACAGAGGCTCCGAATATGACCAACATCTACCTGCTACACCTGAGCCATTTGGAC	255	
241	CGGCTCTCTGCTGATACCTTCCATTTGATGATGATGATGATGATGATGATGATGATGAT	300	CGGCTCTCTGCTGATACCTTCCATTTGATGATGATGATGATGATGATGATGATGATGAT	300	
256	CGGCTCTCTGCTGATACCTTCCATTTGATGATGATGATGATGATGATGATGATGATGAT	315	CGGCTCTCTGCTGATACCTTCCATTTGATGATGATGATGATGATGATGATGATGATGAT	315	
301	TTTGGCCATGAGCATGTGTAAGCTCTCTGAGGGTTTATCACACAGGCTTGTACACGCGAG	360	TTTGGCCATGAGCATGTGTAAGCTCTCTGAGGGTTTATCACACAGGCTTGTACACGCGAG	360	
316	TTTGGCCATGAGCATGTGTAAGCTCTCTGAGGGTTTATCACACAGGCTTGTACACGCGAG	375	TTTGGCCATGAGCATGTGTAAGCTCTCTGAGGGTTTATCACACAGGCTTGTACACGCGAG	375	
361	ATCTTTTTCATATATCTGCTGATACATGACAGTATGATGATGATGATGATGATGATGATGAT	420	ATCTTTTTCATATATCTGCTGATACATGACAGTATGATGATGATGATGATGATGATGATGAT	420	
376	ATCTTTTTCATATATCTGCTGATACATGACAGTATGATGATGATGATGATGATGATGATGAT	435	ATCTTTTTCATATATCTGCTGATACATGACAGTATGATGATGATGATGATGATGATGATGAT	435	
421	GGCCTTTCAGAGCGGAGTCTACTTTTGGTGTATCAGACAGAGTATGATGATGATGATGAT	480	GGCCTTTCAGAGCGGAGTCTACTTTTGGTGTATCAGACAGAGTATGATGATGATGATGAT	480	
436	GGCCTTTCAGAGCGGAGTCTACTTTTGGTGTATCAGACAGAGTATGATGATGATGATGAT	495	GGCCTTTCAGAGCGGAGTCTACTTTTGGTGTATCAGACAGAGTATGATGATGATGATGAT	495	
481	GCAGTGTATGACAGCTCTTCTGATATTTATCTTATGATGATGATGATGATGATGATGATGAT	540	GCAGTGTATGACAGCTCTTCTGATATTTATCTTATGATGATGATGATGATGATGATGATGAT	540	
486	GCAGTGTATGACAGCTCTTCTGATATTTATCTTATGATGATGATGATGATGATGATGATGAT	555	GCAGTGTATGACAGCTCTTCTGATATTTATCTTATGATGATGATGATGATGATGATGATGAT	555	
541	ACTTTTGGAGAGCTTTTACACAGAGGATATGATGATGATGATGATGATGATGATGATGAT	600	ACTTTTGGAGAGCTTTTACACAGAGGATATGATGATGATGATGATGATGATGATGATGAT	600	
556	ACTTTTGGAGAGCTTTTACACAGAGGATATGATGATGATGATGATGATGATGATGATGAT	615	ACTTTTGGAGAGCTTTTACACAGAGGATATGATGATGATGATGATGATGATGATGATGAT	615	
601	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660	
616	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	675	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	675	
661	GGAATCATCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	720	GGAATCATCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	720	
676	GGAATCATCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	735	GGAATCATCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	735	
721	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	775	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	775	
736	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	789	ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	789	

LOCUS	RG182330	884 bp	mrna	linear	EST 21-APR-2001
DEFINITION	RS11196 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.				
ACCESSION	RG182330				
VERSION	RG182330.1	GI:13704017			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 884) Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,P., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,M., and Ducar,M.				
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)				
MEDLINE	21227151				
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scaine@atersys.com High quality sequence stop: 529.				
FEATURES	Location/Qualifiers				
source	1..884 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Athersys RAGE Library" /cell_line="HT1080" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."				
BASE COUNT	191 a 231 c 206 g 256 t				
ORIGIN					
Query Match	66.9%; Score 713; DB 12; Length 884;				
Best Local Similarity	98.3%; Pred. No. 2.3e-202;				
Matches 742; Conservative	0; Mismatches 10; Indels 3; Gaps 2;				
QY	1 ATGACAACCTCACTAGATACAGTTGAGACCTTTGGTACACATCCTACATATATGACGCG 60				
DB	17 ATGACAACCTCACTAGATACAGTTGAGACCTTTGGTACACATCCTACATATATGACGCG 76				
QY	61 GGCCTGCTTTGGAAAAACCTGATACGACAGCAGTGGTGGCCAGTTTGGCCCCGCTG 120				
DB	77 GGCCTGCTTTGGAAAAACCTGATACGACAGCAGTGGTGGCCAGTTTGGCCCCGCTG 136				
QY	121 TACCTCCCTGGTTCACGTGTGGCCTCTTGGCAATGTGGTGGTGGTATGATCCTCATAT 180				
DB	137 TACCTCCCTGGTTCACGTGTGGCCTCTTGGCAATGTGGTGGTGGTATGATCCTCATAT 196				
QY	181 AATATCAGGAGGCTCGGATTTATGACCAACATCTACCTGCTCAACCTGGCCATTTTGGAGC 240				
DB	197 AATATCAGGAGGCTCGGATTTATGACCAACATCTACCTGCTCAACCTGGCCATTTTGGAGC 256				
QY	241 CTGCTCTTCCTGTGATACCTTCATCTTGGATGCCACTATGTACAGGGGGCATTAACGGGTT 300				
DB	257 CTGCTCTTCCTGTGATACCTTCATCTTGGATGCCACTATGTACAGGGGGCATTAACGGGTT 316				
QY	301 TTGGCCATGCGATGTAGCTCCTCTGAGGGTTTATACACAGGCTTTGACAGCGAG 360				
DB	317 TTGGCCATGCGATGTAGCTCCTCTGAGGGTTTATACACAGGCTTTGACAGCGAG 376				
QY	361 ATCTTTTTCATATACCTGCTGACAAACGACAGTACCTGGCCATTTGATCATGCTGTTT 420				
DB	377 ATCTTTTTCATATACCTGCTGACAAACGACAGTACCTGGCCATTTGATCATGCTGTTT 436				

QY 421 GCCCTTCGAGCCCGGAGCTGCTCACTTTGGTGTGTCATCCAGCATGCTGACCTGGGCGCTG 480
|||||
DB 437 GCCCTTCGAGCCCGGAGCTGCTCACTTTGGTGTGTCATCCAGCATGCTGACCTGGGCGCTG 496
QY 481 GGAGGCTGAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
|||||
DB 497 GGAGGCTGAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 541 ACTCTTTGAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
|||||
DB 557 ACTCTTTGAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616
QY 601 CTGAGAGTACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
|||||
DB 617 CTGAGAGTACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 661 GGAATCATCAAAAGCTGTGAGTGGCCCGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGG 720
|||||
DB 677 GGAATCATCAAAAGCTGTGAGTGGCCCGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGG 733
QY 721 ATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
|||||
DB 734 ATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768

RESULT 3
LOCUS B1906283 876 bp mRNA linear EST 16-OCT-2001
DEFINITION 603063222F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212541 5',
mRNA sequence.
ACCESSION B1906283
VERSION B1906283.1 GI:16168946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE NIH-MGC http://mgi.nhl.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11533 row: k column: 06
High quality sequence stop: 800.

FEATURES
Source
Location/Qualifiers
1..876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5212541"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

BASE COUNT 202 a 232 c 193 g 249 t
ORIGIN

Query Match 66.1%; Score 703.8; DB 13; Length 876;

Best Local Similarity 97.2%; Pred. No. 1.3e-199;
Matches 769; Conservative 0; Mismatches 17; Indels 5; Gaps 5;

QY 1 ATGACAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
|||||
DB 79 ATGACAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 138
QY 61 GGCCTGCTGTAAGAAAGCTGATACAGAGCATGATGATGATGATGATGATGATGATGATG 120
|||||
DB 139 GGCCTGCTGTAAGAAAGCTGATACAGAGCATGATGATGATGATGATGATGATGATGATG 198
QY 121 TACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
|||||
DB 199 TACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258
QY 181 AATATACAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
|||||
DB 259 AATATACAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318
QY 241 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
|||||
DB 319 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
QY 301 TTTGGCCATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
|||||
DB 379 TTTGGCCATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 438
QY 361 ATCTTTTATATATATCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||
DB 439 ATCTTTTATATATATCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGAT 498
QY 421 GCCCTTCGAGCCCGGAGCTGCTCACTTTGGTGTGTCATCCAGCATGCTGACCTGGGCGCTG 480
|||||
DB 499 GCCCTTCGAGCCCGGAGCTGCTCACTTTGGTGTGTCATCCAGCATGCTGACCTGGGCGCTG 558
QY 481 GCAATGCTAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
|||||
DB 559 GCAATGCTAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
QY 540 GACTCTTTCAGAGCTCTTTCATCCAGAGATGATGATGATGATGATGATGATGATGATGAT 599
|||||
DB 618 GACTCTTTCAGAGCTCTTTCATCCAGAGATGATGATGATGATGATGATGATGATGATGAT 677
QY 600 TCTGAGAAATGACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
|||||
DB 678 TCTGAGAAATGACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
QY 659 CAGGAATCATCAAAAGCTGTGAGTGGCCCGGAGTAAAGGAGGAGGAGGAGGAGGAGG 717
|||||
DB 738 CAGGAATCATCAAAAGCTGTGAGTGGCCCGGAGTAAAGGAGGAGGAGGAGGAGGAGG 797
QY 718 CTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
|||||
DB 798 CTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 856
QY 778 CTCTCTCTCTC 788
|||||
DB 857 CTCTCTCTCTC 867

RESULT 4
LOCUS BG204024 745 bp mRNA linear EST 21-APR-2001
DEFINITION RSP73417 Athensys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG204024
VERSION BG204024.1 GI:13725711
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,

OY		401	CCATGTCACATCGTGTGGTTGCCCTTGAGGCGCAGACTGCATTGGTGTCACACA	460
Db	241	CCATGTCACACCGCGTGGTGGTGGCGGACGGACCGTCACTTGGTGTCACACA	300	
OY	461	GCATGTCACCTGGGGCGTGGCAGTGCAGACAGCCTTCTGAATTAATCTTCAATAGA	520	
Db	301	GCATCATCATTTTGGGCCCTGGGCATCTGGGCTTCATGCCAAGGCTTAATCTTTCAGA	360	
OY	521	CTGAAGACTTGTGTGAAGAGACTCTTTCAGTGCCTTCTTACCAGAGATACAGATA	580	
Db	361	CCCATGGGAATTCACCTCACACACACTCCAGCTTCACCTTCTCCACAAAGCCATCA	420	
OY	581	GCTGGAGCATTTTCACACACTCGAGAAATGACACTTTCGATGCTGTCCTCCCTGCTG	640	
Db	421	AGTGAAGCTGTTTCAGGCTCTGAAACTGAAACCTTTTGGCGTGTATTCCTTTGTTGG	480	
OY	641	TTCATGCGCATCTGCTACACAGGAATCATCAAACGCTCTGAGTGCCCCAGTAAAAA	700	
Db	481	TTCATGATCATCTGCTACACAGGATTAATAAGATTCCTCTAAGACGACCAATGAGA	540	
OY	701	AGTACAAAGCCATCCGGCTCATTTTTGTGATCATGTGGCGGTGTTTCATTTTCTGAC	760	
Db	541	AATCCAAAGCGTCCGTTGGATTTTTGTGATCATCATCATCATCTTTTCTCTTTGAC	600	
OY	761	CCTAACATGAGCTATCT	820	
Db	601	CCTAACATTTTACTTACTTATTTCTGTTTTCACAGACTTCTGTTTCACCCATGAGT	660	
OY	821	AGCGAGCAAGCATCTGACCTGCTGATCATGCTGTCAGAGAGGTGATGCTACTCTCA	880	
Db	661	AGCAGAGACAGCATTTTGGACCTGCTGTCAGAGAGGTGATGCTACTCTCA	720	
OY	881	GCTGATCAACCCGCTGATCTACGCTTGTGTGAGAGAGGTTCGGGAATACCTGCGC	940	
Db	721	GCTGTCACACCAATGATCTACGCTTGTGTGAGAGGTTCGGGAATACCTGCGC	780	
OY	941	ACTTCTCTCACAGCAGCTTGTGATCATGACCTGTCAGAGAGGTGATGCTACTCTCA	1000	
Db	781	AGTTTCTCACAGCAGCTTGTGATCATGACCTGTCAGAGAGGTGATGCTACTCTCA	839	
OY	1001	AGAACTGGAAGAACACAGCTC	1022	
Db	840	ACAGCGTGGAGAGGCTCAGCTC	861	
RESULT 6				
LOCUS	BC460984	934 bp	mRNA	linear EST 21-APR-2001
DEFINITION	RS43688 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.			
ACCESSION	BC460984			
VERSION	BC460984.1	GI:13749490		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 934)			
	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,			
	Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.			
	, Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,			
	,E., Veloso,N., Kikka,A., Hess,J., Cothren,K., Lo,K., Offenbacher			
	,J., Dancz,J., and Ducar,M.			
TITLE	Creation of genome-wide protein expression libraries using random			
JOURNAL	activation of gene expression			
MEDLINE	Nat. Biotechnol. 19 (5), 440-445 (2001)			
COMMENT	21227151			
	Contact: Scott J. Cain			
	Athersys, Inc.			
	3201 Carnegie Ave, Cleveland, OH 44115, USA			
	Tel: 216 431 9900			
	Fax: 216 361 9396			
	Email: scaine@atersys.com.			

FEATURES	source	Location/Qualifiers
		1. 934
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone_lib="Altersys RAGE Library"
		/cell_line="HT1080"
		/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT	200 a 188 c 237 g 282 t	27 others
ORIGIN		
Query Match	41.1%:	Score 437.2: DB 12: Length 934:
Best Local Similarity	77.7%:	Pred. No. 1.3e-119:
Matches 589:	Conservative 0:	Mismatches 162: Indels 7: Gaps 5
OY	1	ATGCAACCTCACTGATATAGAGACCTTTGGTACACATCTACTATGATGACGTG 60
Db	75	ATGCAACCTCACTGATATAGAGACCTTTGGTACACATCTACTATGATGACGTG 134
OY	61	GGCCTGCTGTGAAAAAGCTGATACAGAGCAGTGATGGCCAGTTGTGCCCGCTG 120
Db	135	GGCCTGCTGTGAAAAAGCTGATACAGAGCAGTGATGGCCAGTTGTGCCCGCTG 194
OY	121	TACTCCCTGGTGTACAGTGGGCCCTTTGGGCATATGGTGGTGTATGATTCCTATA 180
Db	195	TACTCCCTGGTGTACAGTGGGCCCTTTGGGCATATGGTGGTGTATGATTCCTATA 254
OY	181	AAATACAGAGGCTCCGAATTTATGACAACATCTACTGCTCAACCTGGCCATTTTCGAC 240
Db	255	AAATACAGAGGCTCCGAATTTATGACAACATCTACTGCTCAACCTGGCCATTTTCGAC 314
OY	241	CTGCTCTCTCTGTCACCTTCCATTTCTGATGACACATATGTAGGGGGCATTAACGTGTT 300
Db	315	CTGCTCTCTCTGTCACCTTCCATTTCTGATGACACATATGTAGGGGGCATTAACGTGTT 374
OY	301	TTTGGCATGGCATGTATAGCTCCTCTCAGGGTTTTTATCACACAGGCTTGTACACGAG 360
Db	375	TTTGGCATGGCATGTATAGCTCCTCTCAGGGTTTTTATCACACAGGCTTGTACACGAG 434
OY	361	ATCTTTTCAATATCCTGCTGACAAATCGACAGGTAC -CTGGCATTTGTCATGCTGTGTT 419
Db	435	ATCTTTTCAATATCCTGCTGACAAATCGACAGGTAC -CTGGCATTTGTCATGCTGTGTT 494
OY	420	TGCCCTTGGAGCGCGAGACGTGACATTTGGTGTCAACACACATCTGTCACCTGGGGCCT 479
Db	495	TGCCCTTGGAGCGCGAGACGTGACATTTGGTGTATATACCANCAATTTGTAACCTTGAAGACT 554
OY	480	GGCAGTGTAGCAGCTCTTCTCGAATTT -ATCTCTATAGACATGAAAG -TTGTTTGA 536
Db	555	GATCATTTTATGTTTTTTTTGATGAATTTATTTTTTAAATGAGAGGAATATGTTGATGT 614
OY	537	AGAGACTCTTTGACATGCTCTTACCAGAG -GATACAGTATATAGCTGGAGGC -ATTT 593
Db	615	GAAAGCTTTTATGATAGATATGTACTGAGAGGATATCACTAGTCAAGTGAAGTGTATGTT 674
OY	594	CCACACTGTGAAAGACCATCTTCTGTCCTCTCCCTCTGCTGCTGATAGGCATCTG 653
Db	675	TGATCTCTTTGTAATAGAGCCCTTTGGAGTGGCTCTACTCTCTGTATATATGTTACTGT 734
OY	654	CTACACAGGAATCAACAAGCGCTGCTGAGGTGCCCCAGTAAAAAAGATACAGGCCAT 713
Db	735	GGTATTCAGGAATCTGTGCCCCCTCTCCGATGTGCTCTACATCATATATACACAGCTCT 794
OY	714	CCGGCTCAATTTTGTGCATCATGAGCGGTGTTTTCATTT 751
Db	795	CTGGCTGACTTTTGTGTACCGGTGCCGCTGTGCATTT 832

BM917063 1074 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT_6702060 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5483554
 DEFINITION 5', mRNA sequence.
 ACCESSION BM917063
 VERSION BM917063.1 GI:19367442
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1074)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM2011 row: 9 column: 11
 High quality sequence stop: 718.
 Location/Qualifiers

FEATURES

source

1..1074
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5483554"
 /clone_11b="NIH_MGC_106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC library."

BASE COUNT 240 a 282 c 245 g 302 t 5 others
 ORIGIN

Query Match 40.6%; Score 432; DB 14; Length 1074;
 Best Local Similarity 67.4%; Pred. No. 5.2e-118;

Matches 651; Conservative 0; Mismatches 309; Indels 6; Gaps 3;

QY 1 ATGACAACTCTACTAGATGAGTGGACCTTTGGTACACATCTCTATGATGACGTG 60
 11
 DB 27 ATGGAACCTCCAAACACACAGAGACTATGACACGACACAGATTGCTATGGGAT 86
 11
 QY 61 GGGCTCTCTGTGAAAAAGCTGATACAGACGACATGGCCCGCCCGCGAG 120
 1
 DB 87 GCACTCCCTGCCAGAAAGTGAAAGAGGCGCTTTGGGCCCACTGCTGCCCTCTG 146
 11
 QY 121 TACTCCCTGTGTTTCACTGTGGCCTCTTGGCAATGTGGTGGTGTGATGATCTCAT 180
 11
 DB 147 TACTCTGTGTTTGTCAATGTGGCTGTGTTGGAACATCCTGGTGTGCTGGTCTTGG 206
 11
 QY 181 AATATAGAGAGGCTCCGAATTTATGACCAACATTTACCTGCTCACTGGCCATTTGGAG 240
 11
 DB 207 CATATCAAGAGGCTAAATAAATCATGACAGATCTACTCTGCAACCTGGCCATTTCTGAC 266
 11
 QY 241 CTGCTCTCTCTGTGACCTCTCATCTCTGATGATGATGATGATGATGATGATGATG 300
 11
 DB 267 CTGCTCTCTCTGTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 326
 11
 QY 301 TTGGCCATGAGCATGTGTAAGCTCTCTCAGAGGTTTATACACAGAGCTGTGACAGAG 360
 11
 DB 327 TTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 386
 11

QY 361 ATCTTTTCAATATCTCTGTGACAAATGACAGGATACGTCGGCATTTGCTGTTT 420
 11
 DB 387 ATCTTTTCAATATCTCTGTGACAAATGACAGGATACGTCGGCATTTGCTGTTT 446
 11
 QY 421 GCCCTTGAGAGCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 11
 DB 447 GCCTTGAGAGCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 506
 11
 QY 481 GCAGTCTGAGAGCT 540
 11
 DB 507 GCATCTTGAGAGCT 566
 11
 QY 541 ACTCTTGAGAGCT 600
 11
 DB 567 CACACCTGAGAGCT 626
 11
 QY 601 CTGAGATGAGAGCT 660
 11
 DB 627 CTGAACTGAGAGCT 686
 11
 QY 661 GGAATCATCAAAACGCTGCTGAGGTGAGTCCAGTAAATAAAGTACAAAGCCATCCG 720
 11
 DB 687 GGGATATATAAGATTCTGCTAAGACAGCAATGAGAAATAATCAAACTGTCGGTTG 746
 11
 QY 721 ATTTTGTATCATGAGCGGTGTTTTCATTTCTGTGACACCTCATAGTGCCTATCTT 780
 11
 DB 747 ATTTTGTATCATGATGATCATTTCTTTCTTTGAGCCCTCATATTTGATCTACTT 806
 11
 QY 781 CTCTCTCTCAATCAATCCATCTTATTTGAAATGATGAGAGCGGAGAGATCGAGC 840
 11
 DB 807 ATTTCTGTTTCCAAAGCTCTCTGTTCCACCATGATGAGAGCGAGAGATTTGGAGC 866
 11
 QY 841 CTGTCATCTCTGTTG-ACAGAGGTGATTCGCTACTCTCCACTCTG-CATGAAACCGGTGA 898
 11
 DB 867 CTGCGCTGTCAAGTGAAGAGAGTATGCTGCTACAGCACTGCTGTGNTCAACCCAGTGA 926
 11
 QY 899 TCTAGCGCTTTGTT---GAGAGAGGTTCCGAGTACTCTGCGCATCTTCCACAGG 954
 11
 DB 927 TCTACGCCCTTCTTGTGNTGANAAGGTTCCGAAATGATCCGCGCATTTGTTCCAAAG 986
 11
 QY 955 CACTTG 960
 11
 DB 987 GGCTG 992

RESULT 8
 BG460103 881 bp mRNA linear EST 21-APR-2001
 LOCUS BGA60103
 DEFINITION R5142555 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BGA60103
 VERSION BGA60103.1 GI:13748609
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 881)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
 Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J.,
 Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,
 'E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,
 'J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 21227151
 JOURNAL
 MEDLINE
 COMMENT
 CONTACT: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scaine@atersys.com
 Location/Qualifiers

```
source
1. .881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is H1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H1080 under normal circumstances."
BASE COUNT      218 a      245 c      167 g      248 t
ORIGIN
Query Match      38.6%; Score 410.8; DB 12; Length 881;
Best Local Similarity 81.4%; Pred. No. 1.1e-111;
Matches 597; Conservative 0; Mismatches 123; Indels 13; Gaps 10;
QY 8 CCTCAGTATGATGAGACCTTTGGTACACAT-CTTACTATGATGAGTGGGCTG 66
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 23 CCTCAGTATGATGAGACCTTTGGTACACAT-CTTACTATGATGAGTGGGCTG 82
QY 67 CTCT-GTGAAGAGCTGATACAGAGACTGATGCGCCAGTTTGGCCGCTGATCT 125
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 CTCTGTGAAGAGCTGATACAGAGACTGATGCGCCAGTTTGGCCGCTGATCT 142
QY 126 CTGG--TGTTCACTGTGGGCTCTTGG--GCATGTGTGTGTGTGTATGATCTCAT 182
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 143 CTGGCTGTTCACCTGTGGCTCTTGGCGCAATGAGCGGTCGCGATCCTCATAA 202
QY 183 ATACAGAGAGCTCGAATTTGACAAACATCTACCTGCTGACCTGGGCA-TTTGGAGC 241
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 ATACAGAGAGCTCGAATTTGACAAACATCTACCTGCTGACCTGGGCA-TTTGGAGC 262
QY 242 TCGTCTCTCTCGTAC-CCTTCATCTGTGATCCACTATGTGAG-GGGGATTAACCTGGT 299
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 263 TCGTATCTCCACACACCTTCCATTCGATCCACTATGTGAGGGGATTAACCTGGT 322
QY 300 TTTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 356
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 TTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 382
QY 357 CGAGATCTTTTTCATATCTGCTGACAAATGACAGGATGATGATGATGATGATGATGAT 416
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 ATGATCTTTTGTGATCTGCTGACAAATGACAGGATGATGATGATGATGATGATGAT 442
QY 417 GT-TTGGCCCTGAGCCGCGAGCTGATCTTTGGTGTGATCACCAGCATGCTGAG 475
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 GTCCATACCTTCCAGCCGCGAGCTGATCTTTGGTGTGATCACCAGCATGCTGAG 502
QY 476 GCCTGGAGTGTGAGAGCTCTCGAATTTATCTTCTGATGAGAGTGTGTTG 535
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 GCCTGGAGTGTGAGAGCTCTCGAATTTATCTTCTGATGAGAGTGTGTTG 562
QY 536 AAGAGACTTTTTCAGTCTTTTACCAAGAGATACAGATATAGTGTGAGGATTTTC 595
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 AATAGACTCTCTACACTGCTCTTACCGAAGATATACATATATCTGTTGATGATGC 622
QY 596 ACACCTGTGAGATGACCATCTTGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 655
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 ACACCTGTGAGATGACCATCTTGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 682
QY 656 ACAGAGATGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAG 715
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 ACAGAGATGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAG 741
QY 716 GGCTCATTTTGT 728
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 742 GGCTCATTTTGT 754
RESULT 9
BM917763
```

```
LOCUS      BM917763      952 bp      mRNA      linear      EST 12-MAR-2002
DEFINITION AGENCOURT 6614764 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485097
5' mRNA sequence.
ACCESSION  BM917763
VERSION    BM917763.1 GI:19368142
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 952)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLES    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-remail.nih.gov
           Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
           CDNA Library Preparation: Rubin Laboratory
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM2015 row: 9 column: 18
           High quality sequence stop: 590.
FEATURES
source
1. .952
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5485097"
/clone_lib="NIH_MGC_106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Blood; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NIH_MGC Library."
BASE COUNT      225 a      268 c      209 g      250 t
ORIGIN
Query Match      28.4%; Score 302.8; DB 14; Length 952;
Best Local Similarity 67.0%; Pred. No. 2.8e-79;
Matches 445; Conservative 0; Mismatches 217; Indels 2; Gaps 1;
QY 1 ATGACAACCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 ATGAAAACCTTCAAAACACCAAGAGACTATGACACGACCAAGATTTGACTATGGGAT 118
QY 61 GGCGTCTCTGTGAAGAAAGCTGATACAGAGACCTGATGAGGCGCCAGTTTGGCCGCTG 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 119 GCAACCTCGGCGAAGAGGAGAGAGGCGCTTTGGGCCCAACGCTGCGCCCTCTG 178
QY 121 TACTCCCTGCTTCACTGTGGGCTCTTGGCAATGTGTGTGTGTGTGTGTGTGTGT 180
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 179 TACTCTGTGATATTTGTCTATTTGGCTGTGGAACATCTCTGTGTGTGTGTGTGTGT 238
QY 181 AATATACAGAGGCTCGAATTTATGACCAATCTACCTGCTCAACCTGGCATTTGGAG 240
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 239 CAATACAGAGGCTTAATAAATGACAGAGCTACCTCTGAACCTGGCATTTTCTGAC 298
QY 241 CTGCTCTTCCTGCAACCTTTCATCTGATGATGATGATGATGATGATGATGATGATGAT 300
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 299 CTGCTCTTCCTGCAACCTTTCATCTGATGATGATGATGATGATGATGATGATGATGAT 358
QY 301 TTTGGCATGAGATGTGTAAAGCTCTCTCAAGGTTTATATACACAGGCTGTACAGGAG 360
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 359 TTTGGATGATGATGTGTAAAGCTCTCTCAAGGTTTATATACACAGGCTGTACAGGAG 418
QY 361 ATCTTTTATATCTGCTGACATGACAGGATGACAGGATGATGATGATGATGATGATGAT 420
```

```

Db 419 ATCTTTTCATCAATCCGCTGACGATGACAGGACTGCGCAATCGTCCAGCGCGTGT 478
OY 421 GCCCTGACCCGAGACATGACATTTGGTGTATCATCAGCATCGTCCACTGCGGCGTG 480
Db 479 GCCCTGCGGACGAGACGACTTTGGTGTATCATCAGCATCATCATTTGGCGCTG 538
OY 481 GCAGTGTACAGCTCTTCTGTAATTTATCTGTATGAGACTGAGAGTGTGGAAGAG 540
Db 539 GCATCTTGGCTTCATGACGAGGCTTATCTTTCCAAAGCCCAATGGAAATCATCTCAC 598
OY 541 ACTCTTGCAGTGTCTTTTACCCAGAGATACATATATAGCTGGAGGCTTCCACACT 600
Db 599 CACACCTGACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 658
OY 601 CTGAGATGACATCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 658
Db 659 CTGAACTGAACTCTTGGGCTTGTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 718
OY 659 CAGG 662
Db 719 CCGG 722

```

```

RESULT 10
LOCUS BG220110 747 bp mRNA linear EST 21-APR-2001
DEFINITION RST39882 Athersys RAGE library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG220110
VERSION BG220110.1 GI:13746131
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,J., Veloso,N., Kljka,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.

```

```

TITLE Creation of genome-wide protein expression libraries using random
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9800
Fax: 216 361 9596
Email: scain@athersys.com

```

```

FEATURES
source High quality sequence stop: 523.
location/Qualifiers
1..747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

```

```

BASE COUNT 192 a 142 c 183 g 230 t
ORIGIN

```

```

Query Match 27.6%; Score 294; DB 12; Length 747;
Best Local Similarity 96.2%; Pred. No. 1.1e-76;
Matches 333; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
OY 721 ATTTTGCATCATGCGGTGTTTTCATTTCTGACACCTT-ACAATGCGCTATCCT 779
|||||

```

```

Db 747 ATTATGTCATCATGCGGTGTATTCATCTTCTGACACCTCACAATGTGTATCT 688
OY 780 TCTCTCTTCCATCAATCCATCTTATTTGGAATGACTGTGACGGAGCATCTGGA 839
Db 687 TCTCTTTCATCAATCCATCTTATTTGG-ATGACTGTGACGGAGCATCTGGA 629
OY 840 CCTGTCTATGCTGTGACAGAGGTGATGCGCTTCTTCTTCTTCTTCTTCTTCTTCTT 899
Db 628 CCTGTCTATGCTGTGACAGAGGTGATGCGCTTCTTCTTCTTCTTCTTCTTCTTCTT 569
OY 900 CTACGCTTGTGTGAGAGAGGTTCGGAATGACCTGCGGCACTTCTTCTTCTTCTTCTTCTT 959
Db 568 CTACGCTTGTGTGAGAGAGGTTCGGAATGACCTGCGGCACTTCTTCTTCTTCTTCTTCTT 510
OY 960 GCTCATGACCTGGGAGATACATCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1019
Db 509 GCTCATGACCTGGGAGATACATCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 450
OY 1020 CTCTGTCTTCATCCACAGAGCGGGAACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1065
Db 449 CTCTGTCTTCATCCACAGAGCGGGAACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 404

```

```

RESULT 11
LOCUS BF193051 568 bp mRNA linear EST 02-NOV-2000
DEFINITION BF193051 24485 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF193051
VERSION BF193051.1 GI:11076420
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```

REFERENCE
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Legeyrd,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)

```

```

TITLE Single pass sequencing. Bases called and alt.trimmed with phred
JOURNAL v0.980904.e. Vector identified by cross-match with the -minscore 18
COMMENT Contact: Smith rpl
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

```

```

FEATURES
source Single pass sequencing. Bases called and alt.trimmed with phred
location/Qualifiers
1..568
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORN6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
Seq primer: ATTTAGCTGACACTATAG.

```

```

BASE COUNT 126 a 153 c 126 g 163 t
ORIGIN

```

```

Query Match 27.2%; Score 289.6; DB 12; Length 568;
Best Local Similarity 69.4%; Pred. No. 2e-75;
Matches 394; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
OY 380 TGACATCGACAGGTACTGCGCATTTGTCTATGCTGTGTTGCCCTTGAGCCCGGACTG 439
|||||

```



```
|||||
Db 1 TGACATCGACAGTACAGTCCATGTCATCCGCTTGGCCGCGCTCGGACCA 60
440 TCACCTTTGGTGCATCACCAGCATGCTACCTGGGGCTGGCAGTCTAGCAGCTTC 499
61 TCACCTTTGGTGCATCACCAGCATGCTACCTGGGGCTGGCAGTCTAGCAGCTTC 120
500 CTGAATTTATCTCTATGAGACGTAAGAGTTGTTTGAAGAGTCTTTGCGAGTCTTT 559
121 CCGGCTTCACCTTTTCCAAAGACCCAGTGGAGTTCAAGTACCTTCCACCTGACCTTCAT 180
560 ACCCAGAGATACAGTATATAGCTGGAGCATTTCCACACTCTGAGAGTACCATCTTCT 619
181 TTCTCTATGAAAACCTAACAAATTTGGAAAGCTTCCAGGCTTGAATGAAATCATCTTGG 240
620 GTTCGCTTCCTGCTGCTTATGCGCATTCGTACACAGAGATCATAAACGCTGC 679
241 GCGTGGTGGTCTCTATGCTATGCTATGCTATGCTACACAGGATCATAAAGATCTTGC 300
680 TCAGAGTCCCTCAGTAAAAAAGTACAGGCTACCGGCTCATTTTGTATCATGCGCG 739
301 TCAGAGACCAATGAGAAAGTACCAAGCTGCGCTGATTTTGTATCATGATCA 360
740 TGTCTTTTCAATTTCTGACACCTTACAGTGGCTATCTTCTCTTCTTCTATCATCA 799
361 TCTTCTTCTCTTCTTGGAGCTCCCTACATCTGACTGTTTGTCTTCTTCTTCTTCTTCT 420
800 TCTTATTTGGAATGACTGAGCGGAGCAAGATCTGAGACCTGTCATCTGTCAGCAG 859
421 CCTGTTTACCCATCAGTGTAGGACAGACAGCTGGGCTGGGCTGATCAAGAGCAG 480
860 AGGTGATCGGCTACTCCACTGCTGATGATGAACCGGATGATGATGATGATGATGATGAT 919
481 AGGTGATCGGCTACTCCACTGCTGATGATGAACCGGATGATGATGATGATGATGATGAT 540
920 GGTTCGGGAAGTACCTGCGCCACTTCTT 947
541 GATTCCGCAAGTATCTACGCGAGTCTT 568

RESULT 12
B0053936 1167 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT.6874116 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935880
DEFINITION 5', mRNA sequence.
ACCESSION B0053936
VERSION B0053936.1 GI:19813276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1167)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHC2124 row: f column: 09
High quality sequence stop: 493.
Location/Qualifiers
1. 1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5935880"
/clone_lib="NIH_MGC_106"
```

```
/lissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: blood; Vector: pORF7; Site: 1; XhoI: Site 2;
EcoRI: cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene)
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library"
```

BASE COUNT 259 a 331 c 248 g 328 t 1 others

Query Match 26.3%; Score 279.6; DB 14; Length 1167;
Best Local Similarity 69.6%; Pred. No. 2.8e-72;
Matches 406; Conservative 0; Mismatches 175; Indels 2; Gaps 2;

```
QY 37 ACCACATCTCTATGATGATGACGTGGGCTGCTGTGAAAAACCTGATACAGAGACTG 96
|||||
Db 38 ACCACAGATTTGATGATGGGATGCAACCTCGGCAGAGTGAACGAGGCTTT 97
QY 97 ATGGCCAGTTTGTGCCCCCGCTGTACTCCCTGTTCACGTGGGCTCTTGGCAT 156
|||||
Db 98 GGGGCCCACTGCTGCCCTCTGTACTCTGTGTATTTGTCATTGGCCGTGGTGAAC 157
QY 157 GTGGTGGTGTGATGATCCCTATAAATACAGAGGCTCCGATTTATGACCAACATCTAC 216
|||||
Db 158 ATCTGTGTGCTCTGTGCTGTGTGATACAGAGGCTAAAAACATGACCAACATCTAC 217
QY 217 CTGCTCAACCTGCGCATTTTGGACCTGCTCTCTGTCACCTCTCATCTGATTCAC 276
|||||
Db 218 CTCTGACCTGGCCATTTTGTGACCTGCTCTCTGTCACCTGTCACCTGTCGATGAC 277
QY 277 TATGTACGGGGCATACTGCTGTTTGGCCATGCAATGTGTAACTCTCTAGGCTT 336
|||||
Db 278 TACAGTTGAAAGATGATGCTGGGTTTGGATGATGATGATGATGATGATGATGATGAT 337
QY 337 TATCACACAGCTTTGATACAGCGAGATCTTTTATATCTGCTGACAAATGACAGGTAC 396
|||||
Db 338 TATTTACACAGCTTTGATACAGCGAGATCTTTTATATCTGCTGACAAATGACAGGTAC 397
QY 397 CTGCGCATTTGCTCATGCTGTGTTGCCCTTGCAGCCGCACTGCTCATCTTTGGTGTATC 456
|||||
Db 398 CTGCGCATTTGCTCATGCTGTGTTGCCCTTGCAGCCGCACTGCTCATCTTTGGTGTATC 457
QY 457 ACCAGCATGCTACCTGCGGCTGGCAGTGTAGCAGCTCTCTGTAATTTATCTTAT 516
|||||
Db 458 ACCAGCATGCTACCTGCGGCTGGCAGTGTAGCAGCTCTCTGTAATTTATCTTAT 517
QY 517 GAGACGAAAGATGTTTGAAGA-GACTCTTTCGACAGCTCTTACCAAGATACAGT 575
|||||
Db 518 AAGACCCAAATGGAATTCACCTACACACCTTGCACGCTTCTCAGAAAGCCT 577
QY 576 ATATAGCTGAGGC-ATTTCACACTCTGAGATGACATCTT 617
Db 578 ACAAGATGGAACCTGTTTTCANGGCTCTGAAACTGAACCTCTT 620

RESULT 13
BFI93021 533 bp mRNA linear EST 02-NOV-2000
LOCUS BFI93021
DEFINITION 244450 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BFI93021
VERSION BFI93021.1 GI:11076390
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
1 (bases 1 to 533)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.,
and Keele,J.W.
```


